AD_____

Award Number: W81XWH-08-1-0223

TITLE: Vitamin D, Vitamin D Receptor Polymorphisms, and Breast Cancer Aggressiveness in African American and European American Women

PRINCIPAL INVESTIGATOR: Song Yao

CONTRACTING ORGANIZATION: Health Research, Inc.
Roswell Cancer Institute Div.
Buffalo, NY 14263

REPORT DATE: May 2009

TYPE OF REPORT: ANNUAL SUMMARY

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 21-05-2009 | Annual Summary | 21 Apr 2008 – 20 Apr 2009 |

**4. TITLE AND SUBTITLE**
Vitamin D, Vitamin D Receptor Polymorphisms, and Breast Cancer Aggressiveness in African American and European American Women

**5a. CONTRACT NUMBER**
W81XWH-08-1-0223

**5b. GRANT NUMBER**
BC073351

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Song Yao

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Health Research Inc., Roswell Cancer Institute
Elm & Carlton Sts.
Buffalo, NY 14263

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research
And Materiel Command
Fort Detrick, MD 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
In 509 consecutive breast cancer patients, we found the majority of the breast cancer patients were vitamin D insufficient or deficient at the time of diagnosis. In premenopausal women, serum 25-hydroxyvitamin D levels were lower in those with invasive cancer, particularly those with cancer of high aggressive characteristics including high tumor grade, estrogen receptor negative, and triple negative subtype, compared to patients with carcinoma in situ. The estimated odds ratios and 95% confidence intervals associated with high versus low vitamin D levels in premenopausal women were 0.39 (0.17-0.91) of invasive cancer, 0.32 (0.13-0.78) of poorly differentiated cancer, 0.24 (0.08-0.71) of estrogen receptor negative cancer, and 0.14 (0.03-0.60) of triple negative cancer. No similar associations were observed in postmenopausal women. Our data showed that low vitamin D levels were related to high risk of invasive breast cancer, particularly cancer of high aggressive characteristics in premenopausal but not postmenopausal women, indicating that vitamin D may prevent breast cancer progression and reduce high aggressive cancer risk in young women.

**15. SUBJECT TERMS**
Breast cancer, aggressive characteristics, vitamin D, VDR, racial disparity

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | USAMRMC |
| U | U | U | UU | 15 | 19b. TELEPHONE NUMBER *(include area code)* |

**Table of Contents**

**Page**

INTRODUCTION

The proposed research project is to 1) examine serum 25-hydroxy vitamin D levels and vitamin D receptor (VDR) genetic polymorphisms in association with breast cancer aggressive characteristics, and 2) examine the contribution of vitamin D and VDR polymorphisms to breast cancer racial disparity between African-American (AA) and European American (EA) women. The two objectives are addressed in a two-step approach using two different study populations. The first objective will be examined among breast cancer patients enrolled in the DataBank and BioRepository (DBBR) at Roswell Park Cancer Institute; the second objective will be nested in the Women's Circle of Health Study (WCHS), a large scale case-control study with both AA and EA women. By the end of the first year, we have completed data collection, serum vitamin D measurement, and VDR polymorphism genotyping, and are now performing comprehensive data analysis. We have generated some interesting findings of vitamin D levels and breast cancer aggressive characteristics.

BODY

*Task1. To obtain a broad and solid background in breast cancer epidemiology, biology, pathology, treatment, prevention and applied statistics (Month 1-36).*

I have finished all the required didactic course work and exams. I have been attending a series of meetings that are specifically in or highly correlated with breast cancer and epidemiologic analysis. These include weekly department seminar and Dr. Candace Johnson's vitamin D group meeting, biweekly Dr. Christine Ambrosone's group meeting, Work in Progress, Journal Club, and Cancer Prevention Grand Round, and monthly Breast Cancer Working Group and Breast Cancer Group Meeting. I am the coordinator for our departmental journal club and breast cancer group meeting. In addition, when relevant, I continue attending the institutional Faculty Forum, seminars in other programs and luncheons with extramural speakers. With the support from the Predoctoral Award and my mentor, I attended the AACR Molecular Epidemiology Working Group Conference (May 20-23, 2008. Carefree, AZ) and the AACR Science of Cancer Health Disparity Conference (February 3-6, 2009. Carefree, AZ). Upon the completion of another dissertation project that I had been working on, the dissertation committee agreed on my graduation on February 1$^{st}$, 2009. My dissertation title was "Clinical and genetic risk factors for accelerated bone mineral density loss after blood and/or marrow transplantation." With the approval from the DOD, I am continuing my awarded project in the same laboratory under the same mentor Dr. Christine Ambrosone.

*Task2. To learn and practice molecular epidemiologic analysis (Month 1-36).*

I have bee working closely with our genetic statistician Dr. Lara Sucheston on vitamin D receptor (VDR) gene linkage disequilibrium (LD) map construction, tagSNP selection and haplotype analysis. With Dr. Sucheston's guidance, I have completed all the statistical analysis for my dissertation project, and the first aim of my awarded project.

*Task3. To design and conduct breast cancer epidemiologic studies (Month 1-36).*

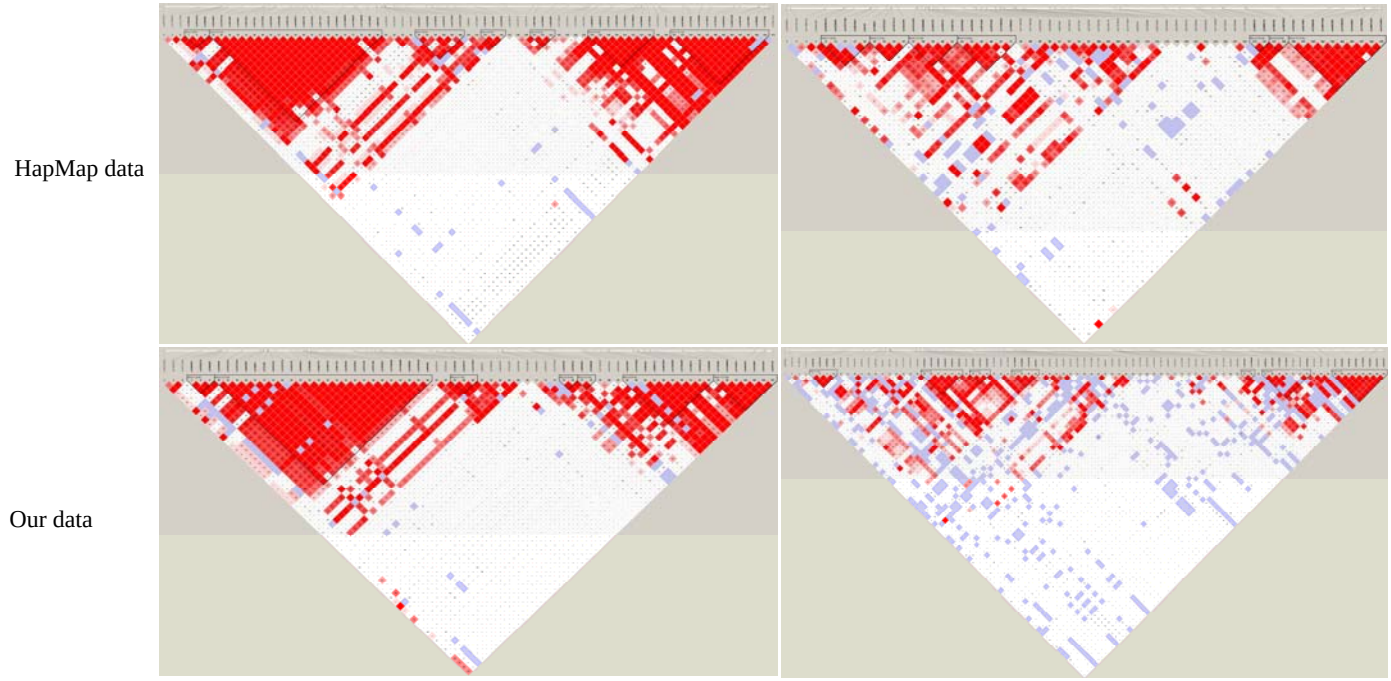
The Specific Aim 1 of my research proposal is to examine the associations between vitamin D levels, *VDR* gene haplotypes and breast cancer aggressiveness (Month 1-15). Human project protection protocol was developed and approved by both the Clinical Research Service (CRS) and Institution Review Board (IRB) at RPCI.

*3-1. To select 70-80 SNPs spanning the extend region of VDR locus from public SNP databases (Month 1).*

VDR gene locates at Chr12q.13.11. It spans a 63.49 kb from 46,521,589 to 46,585,081 (NCBI B35 assembly). To include transcription regulating regions, I extended 15 kb from 5' and 3' end, so the extended genetic region of VDR gene covers a 93.49 kb genetic region from 46,506,589 to 46,590,081. There is no other gene contained in this region. I first surveyed HapMap database for SNPs in this genetic region. Genotype frequency data of this region for both CEU and YRI populations were downloaded from HapMap generic genome browser (Rel 21/phase II Jul 06, on NCBI B35 assembly, dbSNP b125). LD structure of this region was analyzed by HaploView program and shown as **Fig.1**. Consistent with the consensus, LD of YRI population is much low and broken than CEU population. Genotype data were uploaded to Tagger online server for tagSNP selection by

specifying minor allele frequency (MAF)=10%, $r^2$ threshold=1.00, and aggressive multi-marker test. For CEU population, a total of 37 tagSNPs were selected with 100% coverage for 76 eligible SNPs; for YRI population, a total of 56 SNPs were selected with 100% coverage for 63 eligible SNPs. Because the incomplete coverage of HapMap genotyping, I supplemented more SNPs that were not included in the HapMap database by surveying the Genome Variation Server in SeattleSNPs. The same genetic region was interrogated. For CEU population, 99 SNPs with MAF=10% were identified, and 23 SNPs were not in HapMap; for YRI population, 104 SNPs were identified and 41 SNPs were not in HapMap. I also included 7 commonly studied or functional SNPs for CEU and 5 SNPs for YRI population. By combing HapMap tagSNPs, commonly studied SNPs, and additional SNPs from SeattleSNPs, I selected 67 SNPs for CEU population (23% more SNPs than HapMap database), and 102 SNPs for YRI population (39% more SNPs than HapMap database). As a result, there were 122 SNPs for both populations. This is a very high dense SNP panel with a density of 0.77 kb per SNP.

Fig.1 Linkage disequilibrium of VDR gene in African population and European population.




*3-2. To obtain DNA samples of 70 AA and 70 EA healthy controls, and to collaborate with the Microarray and Genomic Facility at RPCI to genotype the selected SNPs (Month 2-4).*

*Genotyping method.* We assembled a preliminary genotyping panel by randomly selecting 70 healthy African American (AA) from the Women's Circle of Health Study (WCHS) and 70 healthy European American (EA) from the Data Bank and Biorepository (DBBR). A hundred nanogram genomic DNA was aliquoted into 96-well plates, dried down, and transferred to the Genomics and Microarray Core, where genotyping was performed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). Ten percent duplicates were included for quality control purpose. MALDI-TOF MS is a very sensitive and high throughput method to detect single nucleotide polymorphisms. A specifically designed oligonucleotide primer is annealed to the DNA immediately one single nucleotide upstream of the polymorphic site. The primer is extended by exactly one single base by a thermo-stable DNA polymerase, and the extension products are desalted, concentrated and subjected to MALDI-TOF for mass analysis. The oligonucleotides are co-precipated with a UV-light absorbing matrix, which is irradiated by a nanosecond laser pulse. The oligonucleotides are ionized by the laser energy and excessive energy is absorbed by the matrix to prevent unwannted DNA fragmentation. The ionized nucleotides are accelerated in an electronic field and then enter the flight tube to be separated according to their mass to charge ratio. The polymorphic site is identified by the mass added by nucleotide onto the primer. Therefore, heterozygous alleles produce two mass-distinct signals that represent the addition of complementary bases to those at the polymorphic site, while homogenous wild or variant alleles produce only one signal. MALDI-TOF MS allows different primers to be added and extended simultaneously in one reaction

as long as there are no cross-reactions between different primers, and thus many different SNPs can be analyzed at the same time.

*Genotyping performance.* The 122 selected SNPs were multiplexed into 7 plexes and primers were designed. Three SNPs failed in the validation by test running in in-house control samples. Among 119 SNPs genotyped, 9 SNPs failed without genotyping calls due to multiple similar genomic sites. Another 2 SNPs failed without known reasons. One trialleleic SNP has good call rate (A+B+E calls) of 52.3% and the MAF deviated significantly from dbSNP data and was thus excluded. Clusters of seven SNPs with good call rate <95% were manually examined and confirmed "C call" as good calls and were included. Among the 107 SNPs included, call rate of each SNP was ≥95%, and for each sample was 100%. Genotype distribution for each SNP was in Hardy-Weinberg Equilibrium. Concordance rate was 100% for all duplicates.

*3-3. To construct VDR gene haplotype block structures using Haploview software, and to select a set of tagSNPs from each block for each population, respectively, using TagSNPs program. This is step will be done with close collaboration with Dr. Sucheston (Month 5-6).*

Currently there are numerous softwares available for genotyping data analysis, but every software has its own requirement for data format. By exploring different approaches, I came across a useful online tool called SNPator (http://bioinformatica.cegen.upf.es/). It is a web-based tool that can be used for genotype data formatting. Spreadsheet of genotype data were uploaded to the server and were converted to Prettybase format. The output data was then analyzed by HaploView for LD structure as shown in Fig.2. As shown, the LD maps of our genotype data were similar to the LD maps of HapMap data. Also, the LD of AA is much lower than that of EA. To select a single set of tagSNPs capturing genetic variations in both AA and EA populations, we used the TAGSter software (http://www.niehs.nih.gov/research/resources/software/tagster/). Genotype data in Prettybase format was converted using a format conversion tool within TAGSter. TagSNPs were then selected by specifying MAF=10% and $r^2$ threshold=0.7. As a result, we selected a set of 61 tagSNPs for both AA and EA populations. The mean maximum $r^2$ for each tagSNP was 0.95.

*3-4. To obtain DNA samples of breast cancer cases from the DBBR, and to genotype selected tagSNPs in all the cases in collaboration with the Microarray and Genomic Facility. Haplotype frequencies will be estimated from unphased genotype frequencies using EM algorithm as implemented in TagSNPs program. (Month 7-10)*

De-identified genomic DNA samples were obtained from 509 newly diagnosed White breast cancer patients who consented to DBBR at RPCI from 2003-2007. 100 ng DNA was aliquoted to 96-well plates, dried down, and transferred to the Genomics Core for genotyping using MALDI-TOF as described above. Because in this aim, only patients self-identified as non-Hispanic White were involved, we further selected 36 tagSNPs from the 61 multipopulation tagSNPs, with an average $r^2$=0.98. Five percent of duplicates were included for quality control. The genotyping performance was good. For each SNP, the call rates was ≥95%; 13 samples with call rates <85% were excluded. The concordance rate between duplicated samples was 100%.

*3-5. To obtain serum samples of the above breast cancer cases and to measure serum 25OHD levels using radioimmunoassay (Month 7-10).*

We obtained de-identified pretreatment serum samples from the same 509 breast cancer patients from the DBBR. We first contacted the Labcorp for 25OHD measurement. Before we sent out the samples, we did a pilot study for assay performance. Serum samples from 15 healthy donors were aliquoted into 4 tubes with 0.5 ml each. The samples were relabeled so that the laboratory personnel were blind about sample matching. Samples were grouped into two batches, and sent to Labcorp in two different times, about 1 week apart. Immunochemiluminometric assay was used on the DiaSorin LIASION automated instrument. We calculated intraclass correlation coefficient (ICC) and coefficient of variation (CV). Between duplicates within each batch, the ICC was 0.63 and 0.88, respectively, and the CV was 20% and 17%, respectively. Between two batches, the ICC was 0.77 and the CV was 19%. Given the unacceptable assay performance, we switched our assay to Heartland Assay, Inc, which was recommended by Dr. Amy Millen. They ran the same assay as Labcorp did, but the performance was much better. We included 5% duplicates in our samples, and the CV was 6.5%, which was acceptable and comparable to what usually reported in literatures.

*3-6. To obtain epidemiologic and clinical data from the DBBR and merge these data with the genotype and haplotype data, as well as the plasma 25OHD levels to generate a master datafile for the subsequent data analysis (Month 11).*

After we obtained genotype data and serum 25OHD data, data managers at DBBR merged the data with clinical and questionnaire data, and identifier that can trace to the patients were removed. The data were then transferred to me for data analysis.

*3-7. To conduct data analysis for Aim 1 (Month 12-15).*

We have completed data analysis for serum 25OHD levels and breast cancer clinical characteristics, and are currently doing analysis for VDR polymorphisms. The patient characteristics were summarized in **Table 1**. The mean age at diagnosis was 57 years, and mean body mass index (BMI) was 28.2 $kg/m^2$. Majority of the patients were postmenopausal (65.4%), and had education more than high school (65.6%). Among the patients, 14.3% had cancer at in situ stage, 53.1% at stage I, 23.5% at stage II, and 9.1% at stage III/IV, and most of the cancer was ductal carcinoma (87.8%). Because of the seasonal variations of vitamin D levels, we classified the time of blood collection into summer/fall season (June-November), and winter/spring season (December-May). Half of the patients had blood collected at summer/fall season, and the other half at winter/spring season. The median 25OHD levels were 55.3 nM, and majority of the breast cancer patients were vitamin D insufficient (41.4%, 50.0-74.9 nM) or deficient (39.7%, <50.0 nM).

Table 1. Descriptive characteristics of patient population

| Characteristics | Cases (n=509) |
|---|---|
| Age at diagnosis, year, mean (SD) | 57.4 (12.4) |
| BMI, $kg/m^2$, mean (SD) | 28.2 (6.3) |
| Menopausal status, n (%) | |
| Premenopause | 176 (34.6) |
| Postmenopause | 333 (65.4) |
| Education, n (%) | |
| Less than high school | 28 (6.6) |
| High school | 118 (27.8) |
| More than high school | 278 (65.6) |
| Family history of breast cancer, n (%) | |
| No | 425 (83.5) |
| Yes | 84 (16.5) |
| AJCC tumor stage, n (%) | |
| In situ | 73 (14.3) |
| I | 269 (53.1) |
| II | 119 (23.5) |
| III & IV | 46 (9.1) |
| Tumor grade, n (%)† | |
| Well differentiated | 40 (9.5) |
| Moderately differentiated | 113 (26.9) |

| | |
|---|---|
| Poorly differentiated | 267 (63.6) |
| ER status, n (%)† | |
| Positive | 332 (77.2) |
| Negative | 98 (22.8) |
| Molecular subtype, n (%)† | |
| Luminal (ER+ and/or PR+) | 337 (78.7) |
| Her2 overexpressing (ER-, PR-, and Her2+) | 29 (6.8) |
| Triple negative (ER-, PR-, and Her2-) | 62 (14.5) |
| Season of blood collection, n (%) | |
| Summer/Fall (June-November) | 256 (50.3) |
| Winter/Spring (December-May) | 253 (49.7) |
| Vitamin D sufficiency, n (%) | |
| Deficiency (<50.0 nM) | 202 (39.7) |
| Insufficiency (50.0-74.9 nM) | 211 (41.4) |
| Sufficiency (≥75.0 nM) | 96 (18.9) |
| Serum 25-hydroxy vitamin D, nM, median (IQR) | 55.3 (40.0-69.9) |

Footnote: † Number of patients and percentage for invasive cases only.

We next examined a number of factors that may affect vitamin D levels and thus may be considered as confounders for adjustment (**Table 2**). Non-parametric Wilcoxon rank test or Kruskal-Wallis test was used for comparisons of 25OHD levels. Premenopausal women had much higher 25OHD levels than postmenopausal women (60.2 nM vs. 53.4 nM, p=0.02), and women with higher BMI had lower 25OHD levels than those with lower BMI (p<0.0001). Patients with higher education also had higher 25OHD levels (p=0.03), and women that were more physically active had higher 25OHD levels (p=0.004). There was no difference in 25OHD levels by family history of breast cancer (p=0.71). Women who had their blood collected in summer/fall season had much higher 25OHD levels than those who had blood collected in winter/spring (62.4 nM vs. 50.0 nM, p<0.0001). We also examined dietary and supplemental vitamin D intake. Generally, patients with higher dietary vitamin D intake and patients taking vitamin D supplement had higher 25OHD levels.

Table 2. Serum 25-hydroxy vitamin D levels by selected covariates

| Characteristics | N | Median (IQR) | P† |
|---|---|---|---|
| Menopause | | | 0.02 |
| Premenopause | 176 | 60.2 (44.3-73.0) | |
| Postmenopause | 333 | 53.4 (38.3-68.3) | |
| BMI category | | | <0.0001 |
| <25.0 kg/m$^2$ | 182 | 63.7 (48.5-77.7) | |
| 25.0-29.9 kg/m$^2$ | 179 | 53.7 (41.5-67.3) | |
| ≥30.0 kg/m$^2$ | 140 | 49.8 (34.3-63.9) | |
| Education | | | 0.03 |

| | | | |
|---|---|---|---|
| Less than high school | 28 | 45.3 (36.1-57.4) | |
| High school | 118 | 53.7 (38.0-69.9) | |
| More than high school | 278 | 56.8 (43.4-70.1) | |
| Outdoor physical activity | | | 0.004 |
| Less active | 65 | 49.4 (35.4-66.3) | |
| Normal | 139 | 53.5 (40.3-67.5) | |
| More active | 211 | 60.1 (44.7-71.2) | |
| Family history of breast cancer | | | 0.71 |
| No | 425 | 55.3 (39.8-69.4) | |
| Yes | 84 | 55.6 (41.4-73.2) | |
| Season of blood collection | | | <0.0001 |
| Summer/Fall | 256 | 62.4 (46.9-76.8) | |
| Winter/Spring | 253 | 50.0 (32.9-62.8) | |
| Dietary vitamin D intake | | | 0.01 |
| Q1 ($\leq$ 30 IU/day) | 108 | 47.7 (29.5-66.4) | |
| Q2 (31-130 IU/day) | 103 | 55.0 (40.7-66.3) | |
| Q3 (131-300 IU/day) | 106 | 58.7 (41.2-73.8) | |
| Q4 (>300 IU/day) | 114 | 56.8 (44.7-70.5) | |
| Supplementary vitamin D intake | | | 0.0004 |
| No | 254 | 52.9 (37.9-66.3) | |
| Yes | 177 | 60.1 (44.7-73.4) | |

Footnote: [†] P-value of Wilcoxon rank test for variables with two levels and Kruskal-Wallis test for variables with more than two levels.

We selected four breast cancer clinical characteristics and compared serum 25OHD levels across each separately (**Table 3**). All comparisons were adjusted by age at diagnosis, BMI, and season of blood collection time in general linear models. Patients with invasive breast cancer had significantly lower vitamin D levels than those with in situ cancer (57.0 nM vs. 55.0 nM, p=0.02). But we found no difference in comparisons of histological grade, ER status, or molecular subtype. Because there was significant difference in vitamin D levels between pre- and postmenopausal women, and etiologic risk factors may be different between the two groups, we stratified our analysis by menopausal status. As shown in **Table 4**, compared to in situ cases, premenopausal breast cancer patients with more aggressive characteristics had lower vitamin D levels; but similar differences were not found in postmenopausal breast cancer patients. We also did comparisons among patients with invasive cancer by using the low aggressive group as reference (well/moderately differentiated, ER+, or luminal, respectively). Premenopausal patients with more aggressive cancer had lower 25OHD levels, though the p-value's were not significant. Compared to patients with luminal subtype (ER+ and/or PR+), those with triple negative disease (ER-, PR-, and Her2-) had significant lower 25OHD levels (56.9 nM vs. 60.2 nM, p=0.04). For similar comparisons in postmenopausal invasive breast cancer cases, there were no differences in vitamin D levels.

Table 3. Serum 25-hydroxy vitamin D levels by selected tumor characteristics

| Tumor characteristics | N | Median (IQR) | P-value[†] |
|---|---|---|---|
| Invasiveness | | | 0.02 |
| In situ | 73 | 57.0 (47.0-77.5) | |
| Invasive | 436 | 55.0 (38.2-69.6) | |
| Histological grade[‡] | | | 0.65 |
| Well/moderately differentiated | 153 | 55.0 (37.3-69.8) | |
| Poorly differentiated | 267 | 55.3 (38.4-69.4) | |
| ER status[‡] | | | 0.55 |
| Positive | 332 | 54.5 (37.2-69.3) | |
| Negative | 98 | 56.4 (41.6-70.4) | |
| Molecular subtype[‡] | | | 0.97 |
| Luminal (ER+ and/or PR+) | 337 | 54.7 (37.1-69.4) | |
| Her2 overexpressing (ER-, PR-, and Her2+) | 29 | 56.7 (41.6-66.9) | |
| Triple negative (ER-, PR-, and Her2-) | 62 | 53.6 (42.1-69.4) | |

Footnote: [†] P-value of Wilcoxon rank test for variables with two levels and Kruskal-Wallis test for variables with more than two levels. [‡] Invasive cases only.

Table 4. Serum 25-hydroxyvitamin D levels by selected tumor characteristics in pre- and postmenopausal women

| Tumor characteristics | Premenopausal | | | | Postmenopausal | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Median (IQR), ng/mL | $P_1$[†] | $P_2$[‡] | N | Median (IQR), ng/mL | $P_1$[†] | $P_2$[‡] |
| Invasiveness | | | 0.004 | - | | | 0.47 | - |
| *in situ* | 29 | 27.5 (21.0-33.8) | | | 44 | 22.0 (17.6-26.2) | | |
| Invasive | 147 | 23.5 (16.0-27.8) | | | 289 | 21.2 (15.0-27.9) | | |
| Tumor grade | | | 0.01 | 0.29 | | | 0.56 | 0.44 |
| Well/moderately differentiated | 43 | 24.2 (17.4-28.0) | | | 110 | 19.7 (14.0-27.9) | | |
| Poorly differentiated | 100 | 23.2 (15.9-27.7) | | | 167 | 21.7 (15.3-27.8) | | |
| ER status | | | 0.009 | 0.29 | | | 0.72 | 0.78 |
| Positive | 108 | 23.9 (15.9-27.7) | | | 224 | 20.6 (14.7-27.6) | | |
| Negative | 38 | 23.0 (18.2-27.8) | | | 60 | 21.6 (16.0-29.3) | | |
| Intrinsic subtype | | | 0.006 | 0.15 | | | 0.86 | 0.93 |
| Luminal | 111 | 24.1 (15.8-27.8) | | | 226 | 20.6 (14.6-27.8) | | |
| Her2 overexpressing | 12 | 24.4 (18.6-29.4) | | | 17 | 22.6 (12.8-24.8) | | |
| Triple negative | 22 | 22.8 (17.4-25.5) | 0.005[§] | 0.04[§] | 40 | 20.8 (16.5-29.3) | 0.72[§] | 0.77[§] |

Footnote: All P-values were adjusted by age of diagnosis, BMI, and season of blood collection time. [†] $P_1$ was derived by using carcinoma *in situ* cases as a reference group. [‡] $P_2$ was derived by using the less aggressive group, i.e., well/moderately differentiated, ER positive, and luminal subtype, respectively, as a reference group. [§] For comparisons between the triple negative subtype and the reference group.

In order to estimate the risk of aggressive breast cancer associated with serum 25OHD levels, we used locally weighted polynomial regression to calculate residuals after controlling for seasonal variations of vitamin D levels (**Fig.2**). For each patient, the residual of vitamin D level is calculated as measured 25OHD levels subtracted by expected levels in the week of blood collection time. Blood collection time adjusted 25OHD levels were calculated as residual plus the population mean. We used the adjusted 25OHD levels to dichotomize patients into high and low vitamin D levels based on the median of in situ patients. Multivariable logistic regression was used to calculate odds ratio (OR) and 95% confidence interval (CI) after controlling for age at diagnosis and BMI. Heterogeneity of associations between pre- and postmenopausal women was calculated by including an interaction term of vitamin D levels times menopausal status. Results were summarized in **Fig.3**. Among premenopausal women, compared to patients with low vitamin D levels, those with high vitamin D levels had lower risk of developing aggressive breast caner. The risk reduction was particularly significant with more aggressive cancer (grade III, ER negative, and triple negative subtype). The ORs of developing more aggressive cancer vs. in situ cancer was 0.32 (95% CI, 0.13-0.78) for grade III cancer, 0.24 (95% CI, 0.08-0.71) for ER negative cancer, and 0.14 (95% CI, 0.03-0.60) for triple negative cancer. However, the ORs in postmenopausal women were all around null. The heterogeneity testings of associations between pre- and postmenopausal women were significant (ER negative risk, p-heterogeneity=0.05; triple negative risk, p-heterogeneity=0.03) or borderline significant (grade III, p-heterogeneity=0.09), indicating differential effects by menopausal status. We repeated our analyses by excluding in situ cancer patients and using low aggressive characteristics groups as reference (**Fig.4**). Although the p-value's were not statistically significant, we still found premenopausal patients with high vitamin D levels had lower risk of developing more aggressive breast cancer: ORs and 95% CI for grade III vs. grade I/II was 0.51 (0.23-1.14); ER negative vs. ER positive was 0.49 (0.20-1.24); and triple negative subtype vs. luminal subtype was 0.30 (0.08-1.11). But the ORs in postmenopausal women were close to 1.00. The heterogeneity also suggested differential effects by menopausal status: p-heterogeneity was 0.14 for grade III vs. I/II, 0.17 for ER negative vs. ER positive, and 0.07 for triple negative subtype vs. luminal subtype.

Fig. 2. Measured and predicted serum 25-hydroxy vitamin D levels by the week in the year of blood collection

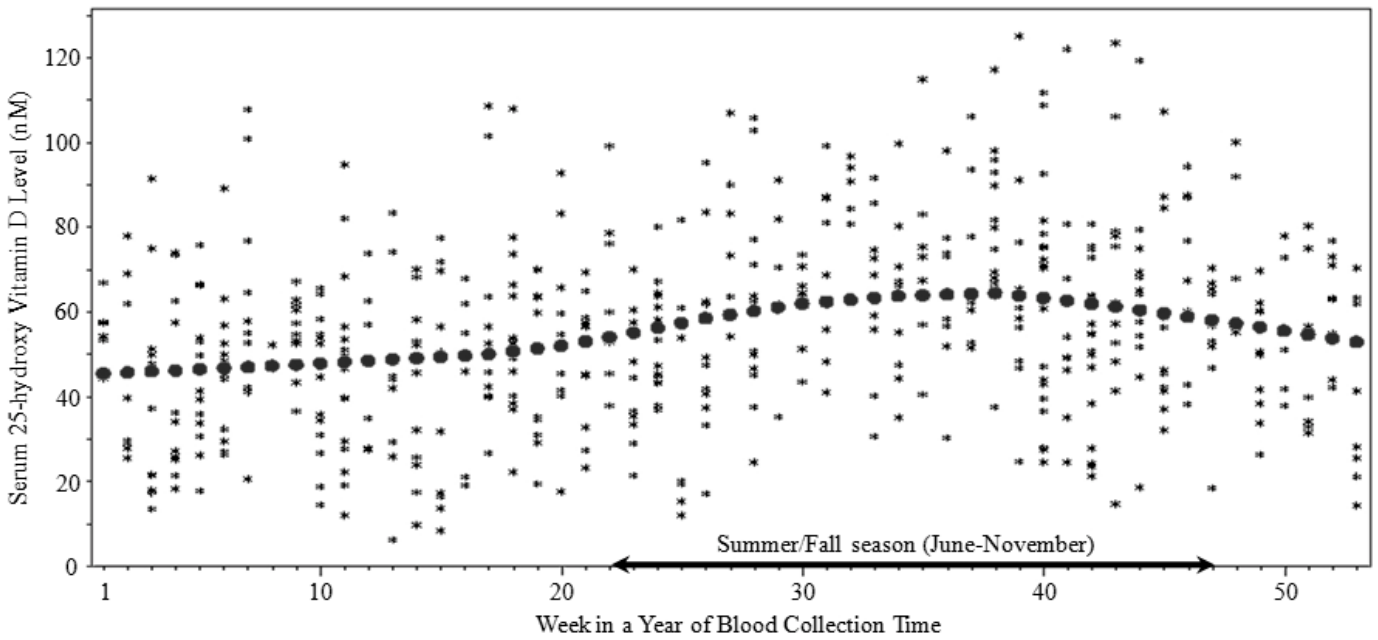
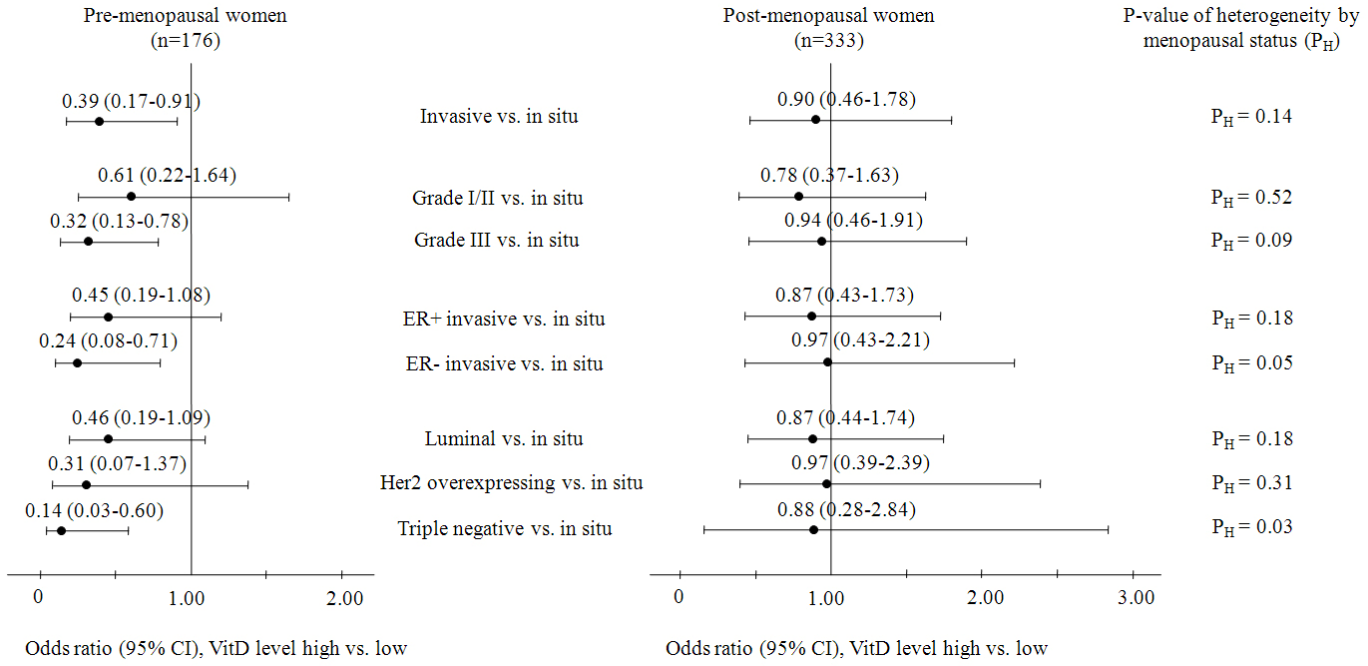
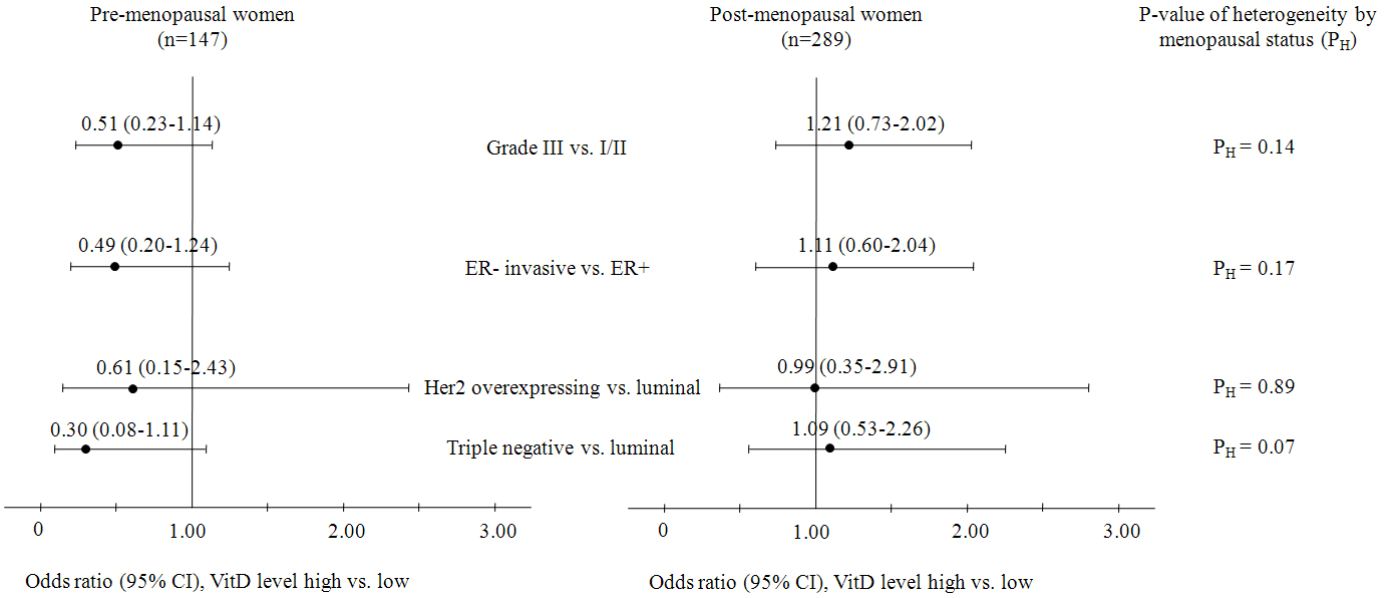


Fig. 3. Odds ratios and 95% confidence intervals of aggressive breast cancer risk associated with serum 25-hydroxy vitamin D levels by menopausal status




Footnote: Serum 25-hydroxy vitamin D levels are stratified into high and low strata based on medians of blood collection time adjusted residuals of in situ cases by menopausal status. Odds ratios and 95% confidence intervals are obtained from logistic regression (invasiveness) or multinomial logistic regression (grade, ER status, and tumor subtype) adjusted for age at diagnosis and BMI.

Fig. 4. Odds ratios and 95% confidence intervals of aggressive breast cancer risk associated with serum 25-hydroxy vitamin D levels by menopausal status in invasive breast cancer patients




Footnote: Serum 25-hydroxy vitamin D levels are stratified into high and low strata based on medians of blood collection time adjusted residuals of in situ cases by menopausal status. Odds ratios and 95% confidence intervals are obtained from logistic regression (grade, ER status) or multinomial logistic regression (tumor subtype) adjusted for age at diagnosis and BMI.

Our results showed for the first time that high vitamin D levels were associated with lower risk of aggressive breast cancer in premenopausal but not in postmenopausal women. The results were consistent with the Nurse's Health Study and Women's Health Study showing that high dietary vitamin D intake was associated with lower overall breast cancer risk in pre- but not in postmenopausal women. We are now conducting analysis of VDR genotype in association with breast cancer aggressive characteristics. As the second aim of the project, we will evaluate vitamin D and VDR genotypes in breast cancer racial disparity between African American and European American women.

KEY RESEARCH AND TRAINING ACCOMPLISHMENTS

- I have finished my predoctoral training and graduated in February 2009. I am continuing the project in the same lab as a postdoctoral fellow with my mentor Dr. Christine Ambrosone.
- We have genotyped 122 single nucleotide polymorphisms (SNP) of vitamin D receptor (VDR) gene in DNA samples from 70 African American (AA) and 70 European American (EA) healthy controls. Linkage disequilibrium structures showed dramatic differences in genetic polymorphisms between the two racial populations. We selected a single set of multipopulation tagSNPs using our own genotyping data to cover genetic variations in both AA and EA populations. They will be genotyping in our study population as a next step.
- We obtained pretreatment serum samples and data from 509 newly diagnosed breast cancer patients. Our analysis showed that in premenopausal women, serum 25-hydroxyvitamin D levels were lower in patients with invasive cancer, particularly those with cancer of high aggressive characteristics including poorly differentiated, estrogen receptor negative, and triple negative subtype. We did not see associations of vitamin D with aggressive breast cancer risk in postmenopausal women.

REPORTALE OUTCOMES

- A manuscript tilted "Serum 25-hydroxyvitamin D levels and breast cancer progression and aggressive characteristics in premenopausal and postmenopausal women: a case-series study" is in preparation to be submitted to the Journal of Clinical Oncology.
- Results generate from this study was used as preliminary data in a DOD Postdoctoral Fellowship Award application titled "Vitamin D local metabolism and breast cancer progression and cancer disparity" submitted in March 2009.

CONCLUSION

To conclude, we found premenopausal women with invasive cancer, particularly those with cancer of high aggressive characteristics including triple negative subtype, had much low serum 25-OHD levels than those with carcinoma *in situ*, indicating that vitamin D may prevent or delay breast cancer progression and reduce risk of breast cancer of high aggressive characteristics. Similar associations were not found in postmenopausal women. The fact that the majority of the breast cancer patients are vitamin D deficient or insufficient at diagnosis confirms the epidemic vitamin D deficiency in the US, especially in breast cancer patients who may benefit from increasing vitamin D levels.

*So what:* Our results show vitamin D may prevent breast progression and reduce the risk of high aggressive breast cancer. If the results are further validated in a prevention trial, young women particularly those at high risk of developing breast cancer shall take vitamin D to prevent breast cancer occurrence and progression.

REFERENCES

None

APPENDICES

None

SUPPORTING DATA

None